Multisensory enhancement of attention depends on whether you are already paying attention

Article (Accepted Version)

Multisensory Enhancement of Attention Depends on Whether You Are Already Paying Attention

Jessica Lunn, Amanda Sjoblom, Jamie Ward, Salvador Soto-Faraco, and Sophie Forster

Abstract

Multisensory stimuli are argued to capture attention more effectively than unisensory stimuli due to their ability to elicit a super-additive neuronal response. However, behavioural evidence for enhanced multisensory attentional capture is mixed. Furthermore, the notion of multisensory enhancement of attention conflicts with findings suggesting that multisensory integration may itself be dependent upon top-down attention. The present research resolves this discrepancy by examining how both endogenous attentional settings and the availability of attentional capacity modulate capture by multisensory stimuli. Across a series of four studies, two measures of attentional capture were used which vary in their reliance on endogenous attention: facilitation and distraction. Perceptual load was additionally manipulated to determine whether multisensory stimuli are still able to capture attention when attention is occupied by a demanding primary task. Multisensory stimuli presented as search targets were consistently detected faster than unisensory stimuli regardless of perceptual load, although they are nevertheless subject to load modulation. In contrast, task irrelevant multisensory stimuli did not cause greater distraction than unisensory stimuli, suggesting that the enhanced attentional status of multisensory stimuli may be mediated by the availability of endogenous attention. Implications for multisensory alerts in practical settings such as driving and aviation are discussed, namely that these may be advantageous during demanding tasks, but may be less suitable to signaling unexpected events.

Keywords: Attention; perceptual load; multisensory integration; attentional capture; exogenous; audiovisual

1. Introduction

Daily life bombards us with an overwhelming amount of sensory input, including sights, sounds, tactile sensations, odours and tastes. Most of them are simply neglected, others instead, summon our attention. Why do certain sensory stimuli attract (or ‘capture’) our attention while others may not be noticed? The types of stimuli argued to capture attention in this way include ‘singletons’ which differ in some unique attribute (e.g. colour) from surrounding items (Theeuwes, 1992), abrupt onsets (Jonides & Yantis, 1988), moving stimuli (Franconeri & Simons, 2003), or events that have motivational relevance or value (Purkis, Lester & Field, 2011; Anderson, Laurent & Yantis, 2013). One type of event that has been proposed to be particularly effective at capturing attention are those which produce correlated stimulation in more than one sensory modality at a time (e.g. Santangelo & Spence, 2007).

Multisensory stimuli are often processed faster or produce stronger responses than unisensory stimuli. According to many studies, this enhanced multisensory response is not merely due to the summed effects of concurrent information, as multisensory stimuli often elicit faster and more accurate responses than would be predicted by additive models of the two unisensory stimuli (Hughes, Reuter-Lorenz, Nozawa & Fendrich, 1994; Colonius & Deiderich, 2002; Molholm et al., 2002; Murray et al., 2004; Laurienti, Kraft, Maldjian, Burgette & Wallace, 2004; Senkowski, Talsma, Hermann & Woldorff, 2005; Talsma, Doty & Woldorff, 2007; Pannunzi et al., 2014; though see Otto & Mamassian, 2012). This has led to the suggestion that multisensory stimuli may also be particularly effective in capturing attention (e.g. Santangelo & Spence, 2007). Whilst this may, under some conditions, be beneficial (i.e. when a multisensory stimulus is of behavioural relevance), it may, on the contrary, be disruptive in other conditions (i.e. by pulling attention away from our current goals). These results have been often been taken to assume that some multisensory integration processes happen prior to, or independent of, the allocation of attention. Contrary to this idea, in the present research we show that whilst multisensory stimuli are particularly effective in capturing attention, this effectiveness is modulated by perceptual load (high load reduces effectiveness in absolute terms, but increases effectiveness relative to that expected from its unisensory parts), and depends on whether or not the stimuli are part of the attentional set (i.e. it is found for targets but not distractors).

1.1. Previous Evidence for Attentional Capture by Multisensory Stimuli

Attentional capture can be measured through both its facilitation effects, whereby a target is identified more rapidly or more accurately when its features capture attention, as well as its distractor interference, whereby a salient but irrelevant distractor disrupts (e.g. slows down) performance from a main task because it summons attention automatically (e.g., Theeuwes, 1992). Facilitatory attentional capture has been found with multisensory stimuli, such as in the ‘pip and pop effect’. The ‘pip and pop effect’ refers to a phenomenon in visual search whereby the presence of an auditory ‘pip’ in time with a target colour change, significantly speeds up search times (produces ‘pop out’) in an otherwise difficult (serial) search task (Van der Burg, Olivers, Bronkhorst & Theeuwes, 2008). It is less clear, however, whether attentional capture by multisensory stimuli can lead to increased distractor interference. For example, employing the widely used response competition flanker measure of distraction, Matusz et al. (2015) found no difference in the level of interference from multisensory audiovisual distractors versus unisensory auditory distractors.

Another paradigm used to test the ability of multisensory stimuli to capture attention is the spatial cuing task (e.g. Posner cueing task; Posner, 1980). In this task, spatial cues are presented shortly prior to imperative targets, either at, or away from, the upcoming target location. These can either facilitate or interfere with target detection depending on whether or not they cue the correct target location. Using this task, two studies have provided evidence suggesting that multisensory cues (both audiovisual and audiotactile) can capture attention more effectively, and therefore produce stronger cueing effects, than unisensory cues (Santangelo & Spence, 2007; Santangelo, Ho & Spence, 2008). However, this multisensory superiority was only found if participants were also performing a demanding central Rapid Serial Visual Presentation (RSVP) task. Under such multi-tasking conditions, unisensory cues (but not multisensory cues) were rendered ineffective. The authors of the original study have since provided further evidence to suggest that the ineffectiveness of the unisensory cues during the simultaneous RSVP task was due to the transients from the central RSVP causing a faster disengagement of attention from the cued location prior to target presentation (Santangelo, Botta, Lupiáñez & Spence, 2009). It is as yet unclear why the central transients did not similarly disrupt multisensory cues, and hence whether the preservation of multisensory cuing effects during the RSVP task reflects enhanced attentional capture, or delayed disengagement. Nevertheless, these findings do appear to reflect some kind of 'special' attentional status for multisensory stimuli. Spence and colleagues have pointed to promising implications regarding the application of multisensory cues during demanding tasks in real life contexts and have found, for example, that a multisensory warning signal in a driving simulator appears to be particularly effective at capturing attention and eliciting faster braking responses in emergency situations (Ho, Reed & Spence, 2007).

The suggestion that multisensory stimuli could be particularly effective in capturing attention during demanding tasks, supported by the studies above, is challenged by work indicating that multisensory integration (of auditory and visual information) may be strongly limited when the stimuli occur away from the focus of attention both in terms of behaviour (e.g., Alsius, Navarra, Campbell and Soto-Faraco, 2005; Pápai & Soto-Faraco, 2017) as well as in brain responses (Senkowski et al, 2005; Talsma et al., 2007; Fernández et al. 2015). Alsius et al., examined attentional influences on the McGurk effect (McGurk & MacDonald, 1976), a multisensory illusion whereby conflicting visual and auditory streams cause the observer to perceive a 'fused' event between the two (for example hearing the syllable 'ba' but seeing someone mouth 'ga' may be perceived as 'da'). Using this illusion, they showed that when participants were involved in a dual task, they were less likely to hear the fused event. This implies that multisensory integration was disturbed by the additional task demands (see also Alsius et al., 2014, for related ERP evidence). One could argue that multisensory processes in complex, linguistic stimuli could be more prone to attention modulation than simple flash and beep events. Yet, Talsma and Woldorff (2005, 2007) demonstrated that non-linear event related potential (ERP) responses, indicating multisensory integration of simple audiovisual stimuli in humans, were observed only in conditions when the stimuli were presented at the location being attended to. Based on this research, it seems important to ascertain under what task conditions increased capture occurs, both in terms of understanding the mechanisms of multisensory attention and to inform potential practical applications.

1.2. Load Theory and Multisensory Stimuli

Load Theory provides a useful theoretical framework predicting the contexts in which attentional capture is more likely to occur. Lavie (1995) proposed Load Theory in order to resolve conflicts between early- and late-selection hypotheses (see Lavie, 2010, for a review). The theory posits that attention works with a limited perceptual capacity, automatically processing stimuli until capacity is depleted. During tasks which involve only low perceptual load, spare capacity remaining after processing relevant information spills over to allow processing of other, less relevant, stimuli. On the other hand, under high perceptual load conditions, all processing capacity must be fully devoted to the relevant task and therefore stimuli irrelevant to the primary task are typically not processed.

Load Theory has been supported by a large body of evidence using various different manipulations of load, and various measures of task-irrelevant processing. Perceptual load manipulations fall largely into two categories: One type of manipulation involves performing the same task in conditions with varying amounts of information – for example, searching for a target letter either when presented alone (low load) or among five other non-target letters (high load). The second type of perceptual load manipulation involves using the same stimuli in both conditions, but changing the task so that it becomes more or less perceptually demanding – for example, searching an RSVP stream for a target defined by either a single feature (low load) or a conjunction of features (high load). A key implication of this framework is that increasing the load of a primary task through these methods reduces behavioural interference from irrelevant distractors (e.g., Lavie & Cox, 1997, Forster & Lavie, 2008), decreases BOLD responses in the visual cortex for irrelevant peripheral stimulation (e.g., Schwartz, Vuilleumier, Hutton, Maravita, Dolan & Driver, 2005), and reduces sensitivity to detect both auditory and visual peripheral stimuli presented in the context of a secondary task (Macdonald & Lavie, 2008; Raveh & Lavie, 2015). Within applied contexts, high visual perceptual load in a driving task has been shown to decrease accuracy in recalling when the sound of braking in a car crash occurred (Murphy & Greene, 2016), and to reduce awareness of both driving-relevant, and driving-irrelevant, visual and auditory stimuli (Murphy & Greene, 2015). It therefore appears that the effects of perceptual load can occur cross-modally in real-life scenarios.

Santangelo and Spence (2007) raised the intriguing possibility that the potential to capture attention by multisensory stimuli may be immune to the effects of perceptual load. This was initially concluded because the effects of unisensory cues, but not multisensory ones, were abolished by a dual task condition. However, as noted above, a subsequent study by the same authors changed their interpretation of these results, concluding that the abolition of unisensory cuing effects was not in fact due to perceptual load, but rather due to the dual task condition involving the presentation of an additional stimulus in between cue and target (Santangelo et al., 2009). As such, the question of whether or not attentional capture by multisensory stimuli is immune to perceptual load, remains unanswered. Such an immunity to load has been found in the past for other classes of stimuli thought to be ‘special’ (particularly effective) in their capacity to grab attention, such as human faces (Lavie, Ro & Russell, 2003).

Santangelo and Spence’s proposal has exciting applied implications, such as the utility of multisensory driving and aviation warning-signals during perceptually demanding activities such as driving down a busy street or landing an aircraft. However, the potentially contradictory findings of cross-modal integration being dependent upon attention (Alsius et al., 2005, Senkowski et al., 2005; Talsma et al., 2007, Pápai & Soto-Faraco, 2017) appear to suggest that, rather than being immune to load effect, multisensory stimuli may not be integrated when conditions demand high levels of attention. If irrelevant multisensory stimuli are not integrated under high load, they would presumably lose the ‘special’ quality that enables them to capture attention so effectively.

1.3. Could the Special Attentional Status of Multisensory Stimuli Itself Depend on Attention?

The evidence discussed above points to a paradoxical situation whereby multisensory stimuli appear to require attention before they acquire the quality that enables them to capture attention. One clue as to how this paradox might be resolved lies in the measures used to demonstrate multisensory attentional capture. As mentioned above, unisensory forms of attentional capture have been widely demonstrated using both facilitation and distraction measures. By contrast, the most convincing behavioural evidence for enhanced multisensory attentional capture involves facilitation effects. Although facilitation effects are widely used as measures of attentional capture, and also most relevant to the applied contexts discussed above, it should be noted that such effects involve stimuli which have already been allocated some top-down attention given that they are part of a search array (e.g. the ‘pip and pop’ effect). As such, it might be that this attentional allocation is sufficient to allow multisensory integration and hence heightened attentional capture. This account could explain the lack of evidence of heightened distractor interference from multisensory stimuli which are not part of the task set – such stimuli would not be allocated sufficient top-down attention to integrate, preventing their enhanced attentional status. The present work examines this possibility by using both facilitation and distraction measures to test for attentional capture by multisensory stimuli.

A second key question raised by the previous literature is whether multisensory attentional capture is, like unisensory attentional capture, dependent upon the availability of perceptual capacity. As discussed above, in unisensory contexts increasing the perceptual load of a primary task has been found to powerfully undermine processing of stimuli, whether these are irrelevant distractors (e.g., Forster & Lavie, 2008) or search targets in a secondary task (e.g., Macdonald & Lavie, 2008). To address whether or not multisensory stimuli are, as has been suggested, immune to these load effects, the current research tested the effects of established manipulations of perceptual load on both multisensory facilitation of secondary target detection (generalising across paradigms and peripheral target salience in Experiments 1, 2 & 3), and interference by multisensory distractors (Experiment 4).

The strongest account of multisensory attentional capture - that multisensory stimuli can capture attention in a purely stimulus driven manner and are immune to any effects of perceptual load – would predict that multisensory stimuli occurring away from a primary task should produce both facilitation and distraction effects, irrespective of perceptual load of that primary task. If on the other hand multisensory enhancement of attentional capture is subject to some form of attentional modulation, this might manifest in two different ways (which are not mutually exclusive). If multisensory attentional capture depends on the allocation of resources regulated by top-down attention, then this would manifest only as facilitation effects and not distraction effects. If multisensory attentional capture depends on the availability of perceptual capacity, it would be eliminated altogether when perceptual load in the primary task is increased.

2. Experiment 1: Serial Visual Search with Unisensory v. Multisensory Peripheral Targets and High v. Low Load Central Task

To address the effects of perceptual load on facilitatory attentional capture by multisensory stimuli, Experiment 1 adapted an established perceptual load manipulation (e.g., Bahrami, Carmel, Walsh, Rees & Lavie, 2007), in which participants search a central RSVP stream for either a single feature (colour, low load) or a conjunction (colour and shape conjunction, high load). Unlike dual versus single task comparisons (e.g. Santangelo & Spence, 2007), our manipulation of the level of load within a task allows us to isolate any influence of load from the effects of single-vs-dual task, on multisensory attentional capture. While performing the central task, participants were also asked to detect peripheral stimuli which could be either multisensory or unisensory. Facilitatory attentional capture in this paradigm would manifest as faster reaction times to multisensory versus unisensory peripheral targets. We should observe a multisensory facilitation at least for the low load task and, an effect of load for unisensory targets. The question is whether load will affect or not the responses to multisensory targets.

2.1 Materials and Methods

2.1.1. Participants

40 participants (26 female) aged between 18 and 35 years (*M* = 23.20, *SD* = 3.68) were recruited at the University of Sussex. All studies were approved by the University of Sussex Sciences & Technology Cross-Schools Research Ethics Committee. 20 participants completed Experiment 1a, and 20 participants completed Experiment 1b. A sample size calculation conducted using G*Power software (Faul et al., 2009) revealed that to detect an effect size of $\eta^2 = .19$ ($\alpha = .05$; $1-\beta = .80$), a sample of 18 participants was required for each Experiment 1a and 1b. The expected effect size was taken from the main effect of cue type in Santangelo and Spence (2007) comparing multisensory and unisensory stimuli. All participants reported normal or corrected-to-normal vision and hearing. Both Bayesian and null hypothesis testing is reported given that the latter is more widely understood, but only the former provides a measure of evidence regarding whether the null or alternative hypothesis is supported by the data (Wagenmakers et al., 2015). Note that Bayesian analysis does not depend on the stopping rule and thus the measure of evidence is valid regardless of stopping rule (see Dienes, 2014; Rouder, 2014). All participants achieved over 75% average accuracy across the experiment.

2.1.2. Stimuli and Procedure

The experiment was programmed and presented using E-prime v2.0, on a 17-inch Dell flat screen, placed 50cm from the participants face, at eye level. Viewing distance was maintained using a chin rest. Loudspeakers, positioned left and right of the screen, were used to present sounds. Each trial began with a central fixation cross presented for 500ms, followed by a stream of nine coloured characters (each subtending $2.3^{o} \times 1.1^{o}$), presented centrally one at a time. Part of the task was to monitor a central stream of characters which were either an S or a 5, and could be coloured red, green, yellow, blue, purple or turquoise. In the high load condition, the target was either a green 5 or yellow S, whereas in low load the target was any red character. Participants reported detection of the target with a foot pedal. Targets appeared as either the 3rd or 6th stimulus in a trial. The timing of presentation was irregular, to increase demand. This was achieved by randomising presentation time of each character (167, 267 or 367ms) with a fixed interstimulus interval (ISI) of 233ms. All stimuli were presented on a light grey background.

In addition to the central task, participants were asked to monitor for peripheral targets which appeared on 50% of trials, presented to the left or right of the central stream. These were presented concurrently with a non-target central stimulus, and therefore did not interfere with responses to the central task. Participants were required to press the left button on a response box if the peripheral target was on the left, and the right button on a response box if it was on the right. Half of these peripheral targets were unisensory, and half were multisensory (i.e. each occurring on 25% of trials). In Experiment 1a, the unisensory target was a black circle of 1.7^{o} diameter (visual only; 100ms), while in Experiment 1b it was a ‘beep’ sound (auditory only; 100ms, 1100Hz). In both experiments, the multisensory target was both the black circle and the ‘beep’ presented together. The unisensory peripheral targets (circle or sound) as well as the multisensory one were presented on either the left or right of the screen (in multisensory targets, the circle and ‘beep’ always occurred at the same side).

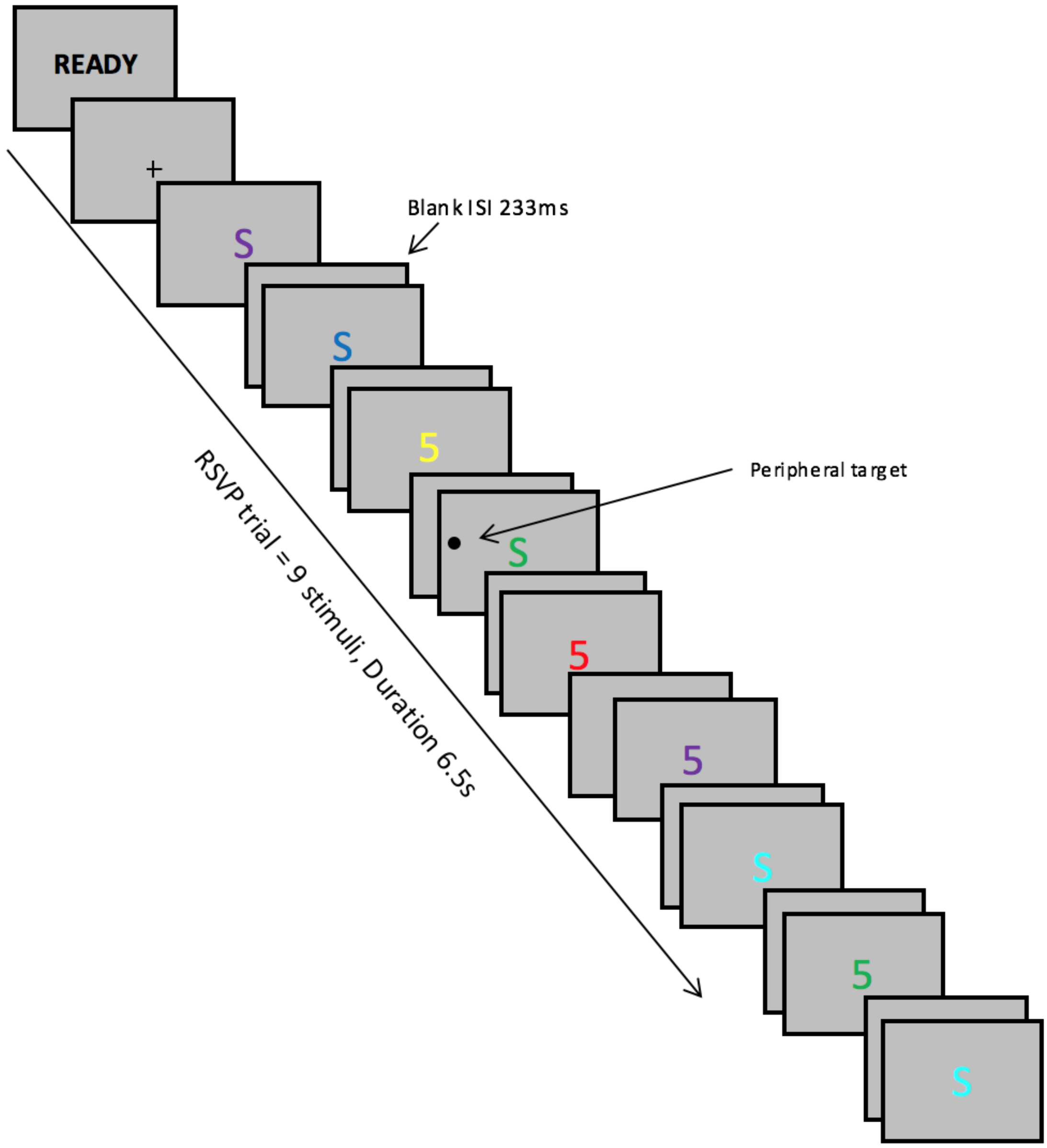

Figure 1: Example RSVP trial, with visual peripheral target

Participants completed two blocks of 144 trials for each load condition, in the order ABBA or BAAB, which was counterbalanced between participants.

2.1.3. Bayesian Analysis

For all tests, *p*-values are reported. Additionally, Bayes factors (B) are reported for all one degree of freedom tests and planned contrasts. Bayes factors (B) were used to assess the strength of the evidence for H1 relative to H0 (Wagenmaker et al., 2015). A B of above 3 is indicative of substantial evidence for H1, whereas a B of below 1/3 indicates substantial evidence for H1, and between these values indicates the data is insensitive (Dienes, 2014).

Bayes factors were calculated using a half-normal distribution, as predictions were all directional, here referred to as $B_{H(0, x)}$, where x is the SD of the distribution. These SDs were based on the results found by Santangelo and Spence (2007) regarding the differences between the distracting effects of peripheral targets of different modalities, and Forster and Lavie (2008) regarding load effects.

2.2. Results and Discussion

Data for all experiments can be downloaded from the Open Science Framework (osf.io/cvy8k).

2.2.1. Reaction time (RT)

Inter-participant average RTs to the central task (correct responses only) were significantly slower under high load (for Experiment 1a M = 753, SD = 100; for Experiment 1b M = 749, SD = 105) than under low load (for Experiment 1a M = 541, SD = 81; for Experiment 1b M = 544, SD = 85; $t(19) = 12.66$, $p < .001$, $B_{H(0,300)} = 5.77\text{x}10^{33}$ for the difference in 1a; $t(19) = 16.11$, $p < .001$, $B_{H(0,300)} = 1.60\text{x}10^{55}$ for the difference in 1b), reflecting the increased demands of the high load task.

Correct RTs to the peripheral targets were entered into a mixed ANOVA, with the within-subjects factors of load (low, high) and peripheral target modality (multisensory, unisensory), and the between subjects factor of experiment number (Experiment 1a and Experiment 1b). This revealed no main effect of experiment number, $p = .534$, and no interactions between experiment number and any of the within-subjects factors, ps > .587. In fact, an identical pattern of results was observed in both Experiments 1a and 1b: Two 2 x 2 within-subject ANOVAs with the factors of load (low, high) and peripheral target modality (multisensory, unisensory) revealed a main effect of load for both Experiment 1a, ($F(1,19) = 38.52$, $p < .001$, $\eta^2 = .67$, $BH_{(0,142)} = 2.40\text{x}10^7$), and Experiment 1b ($F(1,19) = 39.63$, $p < .001$, $\eta^2 = .68$, $B_{H(0,142)} = 3.97\text{x}10^7$). As can be seen in Figure 2, detection of the peripheral targets was slowed in both Experiments 1a and 1b in the high load condition relative to the low load condition, suggesting load modulation.

There was also a main effect of peripheral target modality, both in Experiment 1a ($F(1,19)$ = 68.90, $p < .001$, $\eta^2 = .78$, $B_{H(0,21)} = 3.08\text{x}10^{13}$), and Experiment 1b ($F(1,19) = 18.05$, $p < .001$,

η^2 = .49, $B_{H(0,27)}$ = 1808.52), with faster detection of multisensory targets than both unisensory visual or unisensory auditory targets. However, the critical test was the interaction. There was no significant interaction between load and target modality (Experiment 1a p = .58; Experiment 1b p = .44). Rather, detection of both multisensory and unisensory targets alike was significantly modulated by load, in both experiments (Experiment 1a $t(19) = 5.47$, $p < .001$, $B_{H(0,140)} = 3.64\text{x}10^5$ for multisensory stimuli, $t(19)$ = 5.32, $p < .001$, $B_{H(0,140)} = 1.85\text{x}10^5$ for visual; Experiment 1b $t(19) = 5.29$, $p < .001$, $B_{H(0,140)}$ = $1.33\text{x}10^5$ for multisensory stimuli; $t(19) = 4.92$, $p < .001$, $B_{H(0,140)} = 2.51\text{x}10^4$ for auditory). On the other hand, we note that the detection speed advantage for multisensory stimuli was observed to a similar degree in each of the load conditions, in both experiments (Experiment 1a $t(19) = 8.12$, $p < .001$, $B_{H(0,20)} = 8.28\text{x}10^9$ under high load, $t(19) = 5.68$, $p < .001$, $B_{H(0,20)}$ = $1.34\text{x}10^4$ under low; Experiment 1b $t(19) = 3.45$, $p = .001$, $B_{H(0,27)} = 98$ under high load, $t(19)$ = 4.49, $p < .001$, $B_{H(0,27)} = 5582.35$ under low). Hence, multisensory stimuli did not appear immune to load effects, although their advantage over unisensory stimuli remained across low and high load.

2.2.2. Error

Percentage error rates in the central task were significantly higher under high load (for Experiment 1a M = 14.80, SD = 14.88; for Experiment 1b M = 10.20, SD = 12.01) than under low load (for Experiment 1a M = 5.30, SD = 6.32; for Experiment 1b M = 5.70, SD = 10.31; $t(19) = 3.52$, $p = .001$, $B_{H(0,10)} = 166.95$ for the difference in 1a; $t(19) = 2.18$, $p = .021$, $B_{H(0,10)}$ = 3.91 for the difference in 1b).

Error rates in detection of the peripheral targets were entered into a mixed ANOVA, with the within-subjects factors of load (low, high) and peripheral target modality (multisensory, unisensory), and the between subjects factor of experiment number (Experiment 1a and Experiment 1b). This revealed no main effect of experiment number, p = .810, and no interactions between experiment number and any of the within-subjects factors, ps > .164. In fact, an identical pattern of results was observed in both Experiments 1a and 1b: A 2 x 2 within-subject ANOVA with the factors of load (low, high) and peripheral target modality (multisensory, unisensory visual) on error rates to peripheral targets (Table 1) revealed no significant effect of load, or interaction between load and peripheral target modality, for Experiment 1a or 1b (1a ps > .085; 1b ps > .100). In Experiment 1a there was also no main effect of peripheral stimulus modality found (p = 0.506, $B_{H(0,10)}$ = 0.20), however in Experiment 1b percentage error rates for multisensory stimuli were lower than those for auditory ($F(1,19)$ = 13.71, p = .002, η^2 = .42, $B_{H(0,10)}$ = 18.46). These results show that the RT effects were not due to a speed accuracy trade-off. Error rates were generally very low, thus the advantageous nature of multisensory stimuli is reflected mostly in RTs.

Overall, this experiment demonstrated two key findings. First, our results are consistent with existing evidence of multisensory stimuli enhancement of attentional capture by stimuli which are part of the top-down attentional set. Furthermore, consistent with suggestions regarding the applied utility of multisensory cues during demanding tasks, the multisensory advantage over unisensory stimuli remained regardless of load. However, contrary to previous suggestions, multisensory stimuli did not appear entirely immune to load effects, in that increasing perceptual load in a central task slowed detection of peripheral multisensory targets as much as unisensory ones. As such, processing of multisensory and unisensory stimuli alike appears modulated by the availability of attentional capacity.

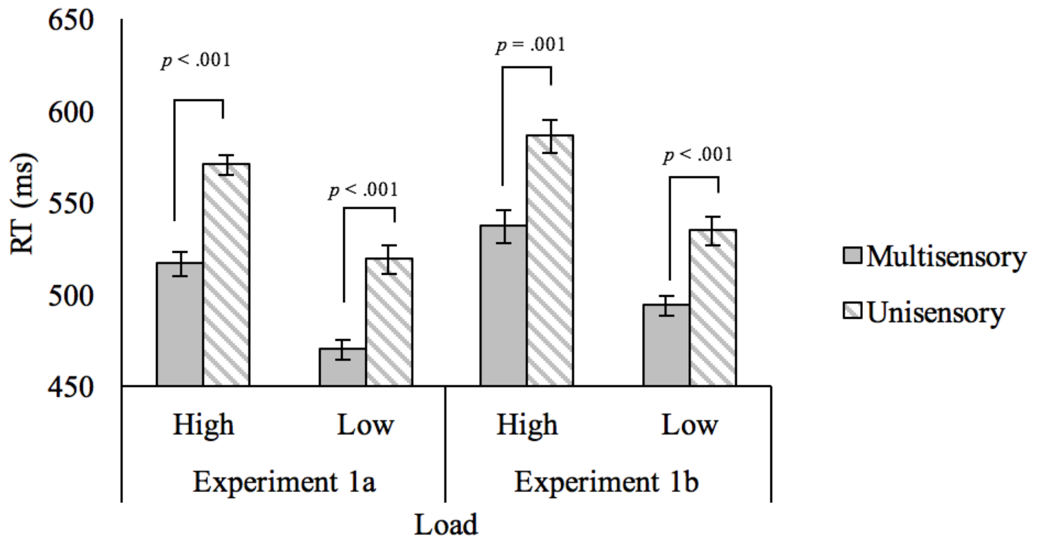

Figure 2: RT (ms) for detection of multisensory and visual only peripheral targets, as a function of load in Experiment 1a, and multisensory and auditory only peripheral targets as a function of load in experiment 1b, error bars show +/- 1 SEM with Cousineau-Morey correction (Cousineau, 2005; Morey, 2008)

		Multisensory	Visual Only	Auditory Only
Experiment 1a	Low Load	6.50 (5.92)	5.90 (6.07)	
	High Load	7.85 (6.33)	9.90 (8.21)	
Experiment 1b	Low Load	5.35 (5.30)		8.00 (5.80)
	High Load	8.05 (7.35)		10.45 (10.66)
Experiment 2	Low Load	4.25 (6.98)	5.38 (7.21)	10.04 (9.25)
	High Load	3.88 (4.30)	6.42 (4.75)	9.46 (8.27)
Experiment 3	Low Load	2.57 (3.65)	5.71 (6.19)	6.32 (6.64)
	High Load	2.14 (3.35)	6.71 (6.66)	7.07 (6.08)

Table 1: Mean percentage error rates (SD in parentheses) as a function of load and target type, across experiments 1-3.

3. Experiment 2: Simultaneous Visual Search with Unisensory v. Multisensory Peripheral Targets and High v. Low Load Central Task

The results of Experiment 1 appear initially consistent with claims that multisensory stimuli can capture attention. These claims derive from evidence that multisensory stimuli elicit super-additive responses. However, data from Experiment 1 could not be tested for non-linear effects, given that the between-subject design did not allow to apply the usual modelling benchmarks (e.g., race model). In the following experiment, we sought to test whether the response to a multisensory stimulus was greater than that which would be predicted by the summed probability of the two unisensory stimuli by testing violations of the race model, which would suggest neural integration of the two sensory stimuli (Miller, 1982; Miller, 1986). In order to be able to calculate race model, auditory and visual unisensory targets were tested within the same experiment, rather than between tasks as in Experiment 1. In addition, Experiment 2 sought to replicate and generalize the findings of Experiment 1 regarding perceptual load effects on multisensory stimuli to another well-established visual search load manipulation in the central task (e.g. Forster & Lavie, 2008).

3.1. Materials and Methods

3.1.1. Participants

26 participants (22 female) aged between 18 and 35 years (*M* = 20.31, *SD* = 2.62) were recruited at the University of Sussex. Two participants were excluded for failing to comply with the instructions. All participants reported normal or corrected-to-normal vision and hearing. The apriori stopping rule for this experiment, and all subsequent experiments, was based on Bayes Factors for the main effect of load and peripheral target type, and all planned comparisons, on reaction time data reaching sensitivity (see Rouder, 2014)[1]. All participants achieved over 75% average accuracy across the experiment.

3.1.2. Stimuli and procedure

The experiment was programmed and presented using E-Prime v2.0, on a 17inch screen. A viewing distance of 57cm was maintained using a chin rest. Loudspeakers positioned on the left and the right side of the screen were used to deliver sounds. Each trial began with a central fixation dot presented for 500ms, followed by a 100ms stimulus display. The stimulus display consisted of six letters (each subtending 0.7^{o} x 0.8^{o}) evenly arranged in an imaginary circle (2.0^{o} radius). On 50% of trials, one of the letters was the target letter which participants were required to search for (X). Participants were required to indicate detection of the target letter by pressing the space bar. In the high load condition, the non-target letters were pseudo-randomly selected from a set of angular letters (H, K, M, V, W, Z, N), whereas in the low load condition the non target letters were all small, placeholder O's (diameter 0.2^{o}). All stimuli were presented on a black background, and all letters were white. The sizing of the stimuli, and the display, was based on previous use of this visual search task (Forster & Lavie, 2007, 2008a, 2008b, 2009, 2014).

On 22.5% of the trials, a peripheral target was presented to either the left or the right of the circular array for letters. On these trials, participants were required to indicate which side of the screen the target was presented on, by pressing one of two keys. These targets could be

[1] This stopping rule was adopted in line with a general change of practice in the lab, in order to determine that any null differences reflect a true no difference between conditions, and are not due to a lack of sensitivity within the data.

unisensory visual, unisensory auditory, or multisensory with equal probability. Peripheral targets were the same as those used in experiments 1a and 1b, and presented for 100ms, with onset at the same time as the central task. They could not appear in the first three trials of each block. All targets were presented to every participant, with the load, central target position, central target identity, peripheral target side, and peripheral target type, fully randomised. Each trial could contain a peripheral target or a central target but not both, or no targets at all. In the latter case (27.5% of trials), no response was required from participants. This prevented the participants from inferring that if there is no peripheral target, that there must be a central one.

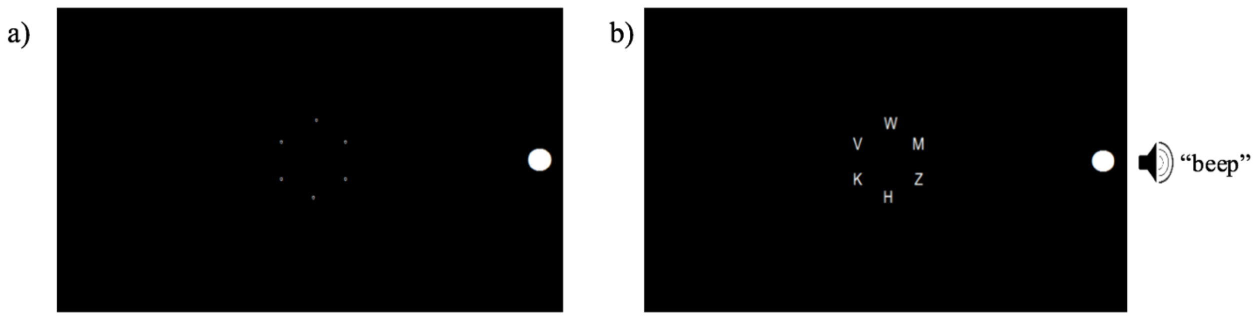

Figure 3: Example stimulus displays: a) low load central letter array with small placeholder letters, with visual peripheral target b) high load central letter array, with multisensory peripheral target

Participants completed three, slowed down, example trials, followed by 12 practice trials, for both high and low load. They then completed four blocks of 80 trials for each load condition, in the order ABBAABBA or BAABBAAB, which was counterbalanced between participants. Participants were instructed to respond as quickly as possible whilst still being accurate (2000ms response deadline).

3.1.3. Bayesian analysis

Bayes factors were calculated using a half-normal distribution, with SDs based on Experiment 1 of this paper.

3.2. Results and Discussion

3.2.1. Reaction time (RT)

Inter-participant average RTs to the central task (correct responses only) were significantly slower under high load (*M* = 634, *SD* = 117) than under low load (*M* = 496, *SD* = 82), $t(23) = 8.18$, $p < .001$, $B_{H(0,212)} = 4.7 \times 10^9$, indicating that the high load task was more demanding.

Correct responses to peripheral targets were entered into a 2 x 3 within-subject ANOVA with the factors of load (low, high) and peripheral stimulus modality (multisensory, unisensory visual, unisensory auditory; Figure 5). As in Experiment 1, the results for peripheral targets revealed main effects of load, $F(1,23) = 23.94$, $p < .001$, $\eta^2 = .51$, $B_{H(0,46)} = 2.36 \times 10^9$, and stimulus modality, $F(2, 46) = 106.45$, $p < .001$, $\eta^2 = .82$, the latter reflecting faster RTs to multisensory peripheral targets compared to either visual only ($B_{H(0,52)} = 6.77 \times 10^6$) or auditory only ($B_{H(0,44)} = 8.68 \times 10^{53}$).

In contrast to the previous experiment, these main effects were qualified by a significant interaction between load and peripheral stimulus modality, $F(2,46) = 13.00$, $p < .001$, $\eta^2 = .36$). This reflected that responses to multisensory targets were modulated by load to a lesser extent than responses to visual targets ($t(23) = 3.90$, $p < .001$, $B_{H(0,120)} = 580.84$), but to an equivalent extent than auditory targets (p = .940, $B_{H(0,146)} = .07$). As in Experiment 1, detection of peripheral targets was slower under high versus low load conditions regardless of sensory modality (multisensory ($t(23) = 3.54$, $p < .001$, $B_{H(0,82)} = 176.73$), visual only ($t(23) = 8.88$, $p < .001$, $B_{H(0,52)} = 9.83 \times 10^{15}$ and, auditory only ($t(23) = 1.88$, $p = .036$, $B_{H(0,52)} = 3.01$). A detection speed advantage for multisensory targets over both types of unisensory target was also observed under conditions of both high load ($t(23) = 10.64$, $p < .001$, $B_{H(0,40)} = 1.05 \times 10^{22}$; $t(23) = 6.24$, $p < .001$ $B_{H(0,54)} = 5.10 \times 10^7$, for multisensory compared with auditory and visual, respectively) and low load ($t(23) = 12.28$, $p < .001$, $B_{H(0,50)} = 2.08 \times 10^{30}$; $t(23) = 2.56$, $p = .009$, $B_{H(0,50)} = 7.83$, for multisensory compared with auditory and visual, respectively).

In order to further determine whether multisensory stimuli provide a detection speed advantage consistent with integration, we used the race model (Miller's inequality, Miller

1982). The race model allows to investigate whether the reaction times in the multisensory condition exceed the statistical facilitation predicted by probability summation based on two independent unisensory signals. In this model, a theoretical cumulative density function (CDF) is calculated based on the reaction time CDFs of each of the two unimodal stimulus types - F_x and F_y - and the redundant-stimulus, or multisensory, condition, F_z. The race model inequality

$$F_z(t) \leq F_x(t) + F_y(t),\ t > 0,$$

for every value of t. Where the empirical CDF towards multisensory stimuli is greater than the theoretical CDFs based on the two unisensory components (tested using paired sample t-tests), the reaction time advantage can be assumed to be caused by integrative effects. Analyses were carried out using the RMITest software, which applies the algorithm in Ulrich, Miller and Schröter (2007).

The results showed that under low load, whilst the reaction time towards multisensory stimuli tends to be faster than the race model bound for the lowest (fastest) percentiles of the reaction time distribution, this does not reach statistical significance ($p > .05$). However, under high load, reaction time to multisensory stimuli is significantly faster than the race model bound for three of the fastest percentiles (Figure 4). This supports the assumption that the detection time advantage under high load might result from cross-modal integration.

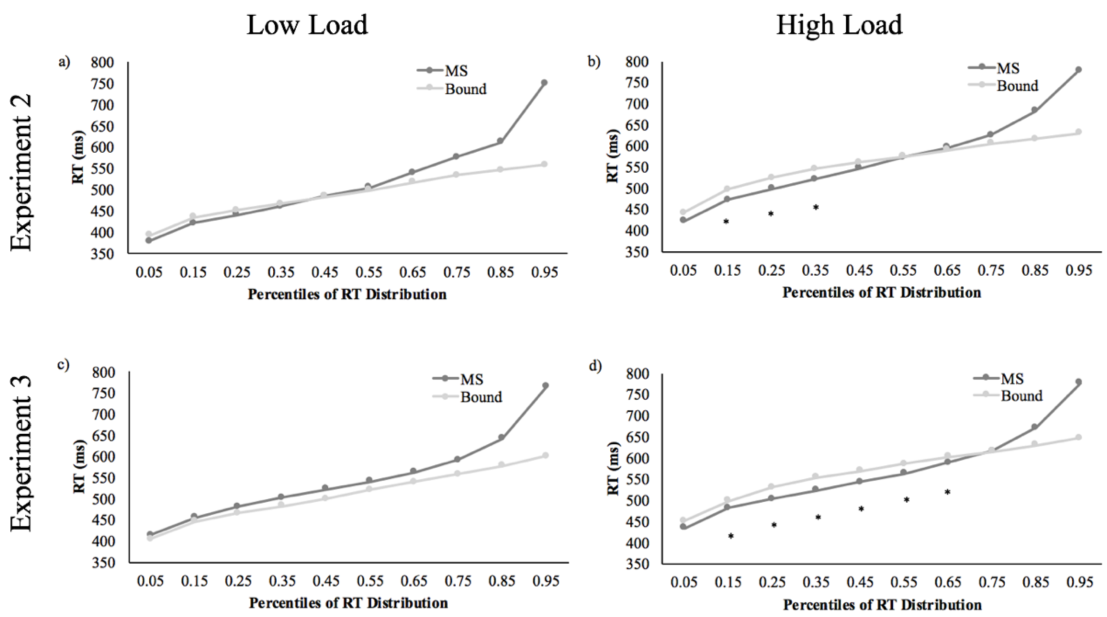

Figure 4: Cumulative probability distribution of reaction times for detection of multisensory (MS) peripheral stimuli, with race model bound for the two unisensory peripheral stimuli predicted by RMITest for a) Experiment 2 low central perceptual load, b) Experiment 2 high central perceptual load, c) Experiment 3 low central perceptual load, d) Experiment 4 high central perceptual load, asterisks refer to where race model inequality was significantly violated, based on Ulrich et al., (2007) algorithm

3.2.2. Error

Percentage error rates in the central task were significantly higher under high load (*M* = 11.75, *SD* = 10.60) than under low load (*M* = 6.42, *SD* = 3.27), *t*(23) = 2.41, *p* = .012, $B_{H(0,7)}$ = 8.85.

Percentage error rates in peripheral stimuli detection were then entered into a 2 x 3 within-subject ANOVA with the factors of load (low, high) and peripheral stimulus modality (multisensory, unisensory visual, unisensory auditory), revealing a main effect of peripheral stimulus modality, $F(1.47, 33.77) = 15.32$, $p < .001$, $\eta^2 = .40$; Table 1. Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) =$ .64, $p = .007$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε =

.73). However, there was no main effect of load (p = .979, $B_{H(0,7)}$ = 0.13), nor a significant interaction between load and peripheral stimulus modality, (p = .668).

Under conditions of high perceptual load, error rates for multisensory stimuli (M = 3.88, SD = 4.30) were significantly lower than those for visual stimuli (M = 6.42, SD = 4.75), t(23) = 2.88, p = .004, $B_{H(0,3)}$ = 26.00, and for auditory stimuli (M = 9.46, SD = 8.27), t(23) = 3.15, p = .002, $B_{H(0,3)}$ = 40.41. This effect was also seen under conditions of low perceptual load for auditory stimuli; error rates for multisensory stimuli (M = 4.25, SD = 6.98) were significantly lower than those for auditory stimuli (M = 10.04, SD = 9.25), t(23) = 3.39, p < .001, $B_{H(0,3)}$ = 74.84, however there was no significant difference in error rates for detection of multisensory and visual peripheral targets (M = 5.38, SD =7.21), (p = .167, $B_{H(0,3)}$ = .93) under low load.

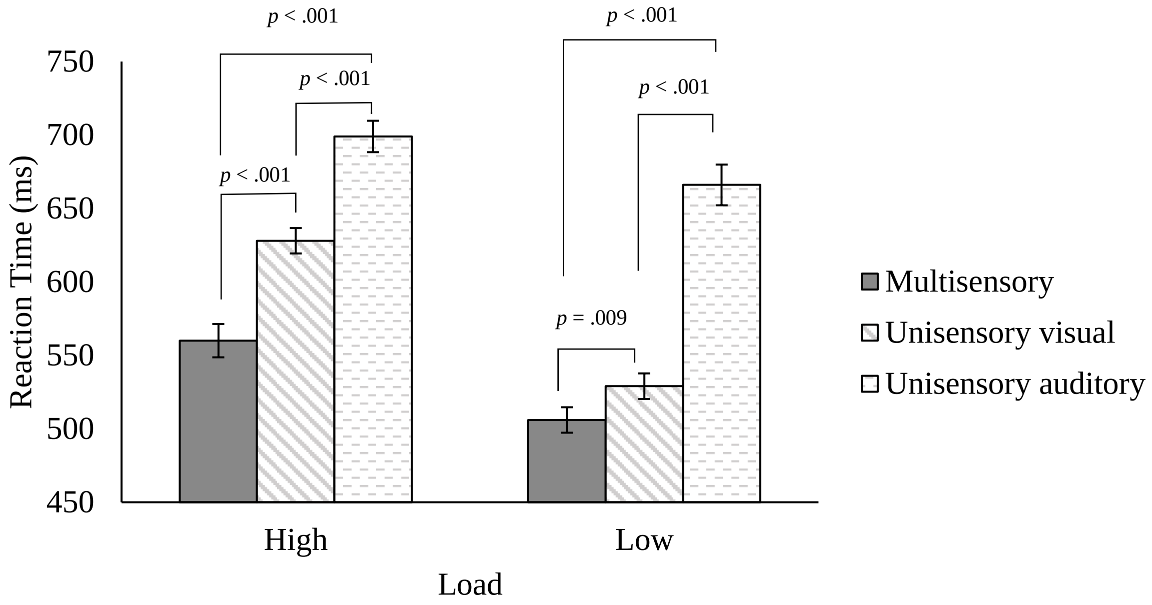

Figure 5: Reaction time (ms) of detection of peripheral targets in different modalities, as a function of load, in Experiment 2. Error bars show +/- 1 SEM with Cousineau-Morey correction (Cousineau, 2005; Morey, 2008)

In summary, Experiment 2 replicates the two key findings of Experiment 1: multisensory stimuli are affected by load modulations, and multisensory stimuli show a detection advantage above unisensory events which is preserved across load manipulations. Again, these results are consistent with an enhancement of attentional capture by stimuli which have been allocated some degree of top-down attention. Unlike Experiment 1, Experiment 2 also found that the reaction time advantage for multisensory stimuli was more pronounced under the high load condition versus the low load condition, with RTs demonstrating that multisensory stimuli were modulated by load to a lesser extent than visual stimuli, and the race model only producing evidence consistent with integration in the high load condition. This raises the intriguing possibility that, while not entirely immune to any effect of perceptual load, multisensory stimuli might be somewhat more resistant to these effects than unisensory events.

4. Experiment 3: Salient Unisensory v. Multisensory Peripheral Targets and High v. Low Load Central Task

Experiments 1 and 2 suggest an advantage for multisensory stimuli over unisensory stimuli. We note, however, that the unisensory stimuli used in these experiments were of relatively low salience, which may have encouraged the multisensory advantage: Following the principle of inverse effectiveness (e.g., Merideth & Stein, 1986), initially proposed for single neuron responses, the weaker the responses to the individual unisensory stimuli, the more likely super-additive responses are to occur if they are presented together as a multisensory event. A remaining question is therefore whether facilitation of attentional capture by multisensory stimuli is limited to low salience stimuli. This appears particularly important given applied suggestions regarding the use of multisensory alerts: in a real world scenario, a unisensory stimulus that is hard to detect would not be reasonably used as an alert or warning signal. Here we addressed whether multisensory stimuli would be capable of facilitating detection above and beyond stimuli highly salient in one unisensory domain, instead of the low saliency events used in Experiments 1 and 2. To test this Experiment 3 repeated the paradigm of Experiment 2, using peripheral target stimuli that are larger, more colourful, meaningful and familiar so that the visual unisensory stimuli would be highly salient relevant to the central task.

4.1. Materials and Methods

4.1.1. Participants

28 participants (22 female) aged between 18 and 33 years (M = 21.04, SD = 2.92) were recruited at the University of Sussex. Participants were recruited until all Bayes Factors for the main effects of load and peripheral target type on reaction time data reached sensitivity.

All participants reported normal or corrected-to-normal vision and hearing. Participants either gained course credits, or were paid, to take part. All participants achieved over 75% average accuracy across the experiment.

4.1.2. Stimuli and procedure

The stimuli and procedure were identical to Experiment 2, with the exception of the identity of the peripheral target stimuli. Visual peripheral targets consisted of a photograph of an animal, randomly selected from six possible images (dog, cat, pig, horse, cow). These were presented in full colour with a black background, subtending 5.0° to 7.5° vertically, by 6.0° to 7.0° horizontally, between 2.5° and 3.0° edge-to-edge from the nearest circle letter. The auditory peripheral targets consisted of the sound each of the six animals makes, played from one of the speakers at the side of the screen (600-1120ms). The multisensory targets were both the animal image and sound, presented on the same side.

Participants completed three slow example trials, followed by 12 practice trials, for both high and low load. They then completed four blocks of 81 trials for each load, in the order ABBAABBA or BAABBAAB, which was counterbalanced between participants. Participants were instructed to respond as quickly as possible whilst still being accurate. They had 2000ms make a response, a short beep indicated where this had been incorrect.

4.1.3. Bayesian analysis

Bayes factors were calculated using a half-normal distribution, with SDs based on Experiment 2 of this paper.

4.2. Results

4.2.1. Reaction time (RT)

RTs to the central task (correct responses only) were significantly slower under high load (M = 635.60, SD = 69.91) than under low load (M = 506.40, SD = 61.28), $t(27) = 12.72$, $p < .001$, $B_{H(0,138)} = 1.22 \times 10^{34}$, indicating that the high load task was more demanding.

Correct responses to peripheral targets were entered into a 2 x 3 within-subject ANOVA with the factors of load (low, high) and peripheral stimulus modality (multisensory, unisensory visual, unisensory auditory; Figure 6). As in previous experiments, this revealed main effects of load, $F(1,27) = 30.59$, $p < .001$, $\eta^2 = .53$, $B_{H(0,62)} = 6.47\text{x}10^9$, and peripheral stimulus modality, $F(1.50, 40.57) = 317.22$, $p < .001$, $\eta^2 = .92$ (Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = .67$, $p = .005$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity, $\varepsilon = .75$) with RTs to multisensory peripheral targets being faster than to either visual only ($t(27) = 4.72$, $p < .001$, $B_{H(0,46)} = 1.55\text{x}10^4$) or auditory only ($t(27) = 24.25$, $p < .001$, $B_{H(0,148)} = 2.69\text{x}10^{126}$). As in Experiment 2, this was qualified by a significant interaction between load and peripheral stimulus modality, $F(2,54) = 5.45$, $p = .007$, $\eta^2 = .17$. Responses to multisensory targets were modulated by load to a lesser extent than responses to visual targets ($t(27) = 2.96$, $p = .003$, $B_{H(0,45)} = 30.50$), but not less than auditory targets ($p = .552$, $B_{H(0,45)} = 0.26$).

As in both previous experiments, detection of all three types of peripheral target was slower under high versus low load (multisensory ($t(27) = 4.02$, $p < .001$, $B_{H(0,52)} = 829.07$), visual only ($t(27) = 5.69$, $p < .001$, $B_{H(0,100)} = 2.07\text{x}10^6$) and auditory only ($t(27) = 2.80$, $p = .005$, $B_{H(0,32)} = 21.39$). The detection speed advantage for multisensory targets over auditory only was observed under both high ($t(27) = 15.18$, $p < .001$, $B_{H(0,138)} = 7.32\text{x}10^{48}$) and low ($t(27) = 19.62$, $p < .001$, $B_{H(0,160)} = 2.54\text{x}10^{82}$) load conditions. The advantage over visual only targets was significant under high load ($t(27) = 5.78$, $p < .001$ $B_{H(0,68)} = 3.23\text{x}10^6$), however no sensitive evidence was obtained under low load ($p = .082$, $B_{H(0,22)} = 1.54$).

RMITest software (Ulrich et al., 2007) was again employed to test for violation of the race model inequality. Similar to Experiment 2, the results showed that under low load there were no significant violations of the race model, whereas under high load reaction time to multisensory stimuli was significantly faster than that which would be predicted by the race model across most of the fastest percentiles, indicating again that integration is occurring and resulting in the faster detection times (Figure 4).

4.2.2. Error

Percentage error rates in the central task were significantly higher under high load (*M* = 10.38, *SD* = 8.55) than under low load (*M* = 7.00, *SD* = 4.12), $t(27) = 2.67$, $p = .006$, $B_{H(0,6)} = 22.31$.

A 2 x 3 within-subject ANOVA with the factors of load (low, high) and peripheral stimulus modality (multisensory, unisensory visual, unisensory auditory) on percentage error rates of peripheral stimuli detection was conducted (Table 1). As in Experiment 2 there was a main effect of peripheral stimulus modality, $F(1.64,44.33) = 13.10$, $p < .001$, $\eta^2 = .33$. Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = .78$, $p = .041$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .82$). There was no main effect of load found ($p = .680$, $B_{H(0,6)} = 0.22$), nor a significant interaction between load and peripheral stimulus modality, ($p = .702$).

Under conditions of high perceptual load, error rates for multisensory stimuli (*M* = 2.14, *SD* = 3.35) were significantly lower than those for visual stimuli (*M* = 6.71, *SD* = 6.66), $t(27) = 3.79$, $p < .001$, $B_{H(0,3)} = 345.15$, and for auditory stimuli (*M* = 7.07, *SD* = 6.08), $t(27) = 4.22$, $p < .001$, $B_{H(0,6)} = 1982.36$. These effects were also seen under conditions of low perceptual load; error rates for multisensory stimuli (*M* = 2.57, *SD* = 3.65) were significantly lower than those for visual stimuli (*M* = 5.17, *SD* = 6.19), $t(27) = 2.78$, $p = .005$, $B_{H(0,3)} = 8.81$, and for auditory stimuli (*M* = 6.32, *SD* = 6.64), $t(27) = 2.73$, $p = .006$, $B_{H(0,6)} = 15.61$.

Thus, this result further indicates an advantage for detection of multisensory stimuli. Even when engaged in a perceptually demanding central task, participants were able to detect the spatial location of a multisensory target with more accuracy than either visual or auditory alone.

In summary, the results of Experiment 3 demonstrate that multisensory stimuli can enhance facilitatory attentional capture even for stimuli that are already highly salient. Consistent with the previous experiments, the results do not support the strongest claim of multisensory immunity to effects of perceptual load insofar as the detection of multisensory targets was slowed down under high load. However, as in Experiment 2 both multisensory and auditory targets were modulated by load to a lesser extent than visual unisensory targets. Furthermore, also as in Experiment 2, the race model analysis only showed sensitive evidence of an integration-facilitated detection advantage in the high load condition. Taken together these findings support the notion that multisensory integration can lead to benefits in detecting searched for stimuli, which may be particularly apparent during more demanding tasks.

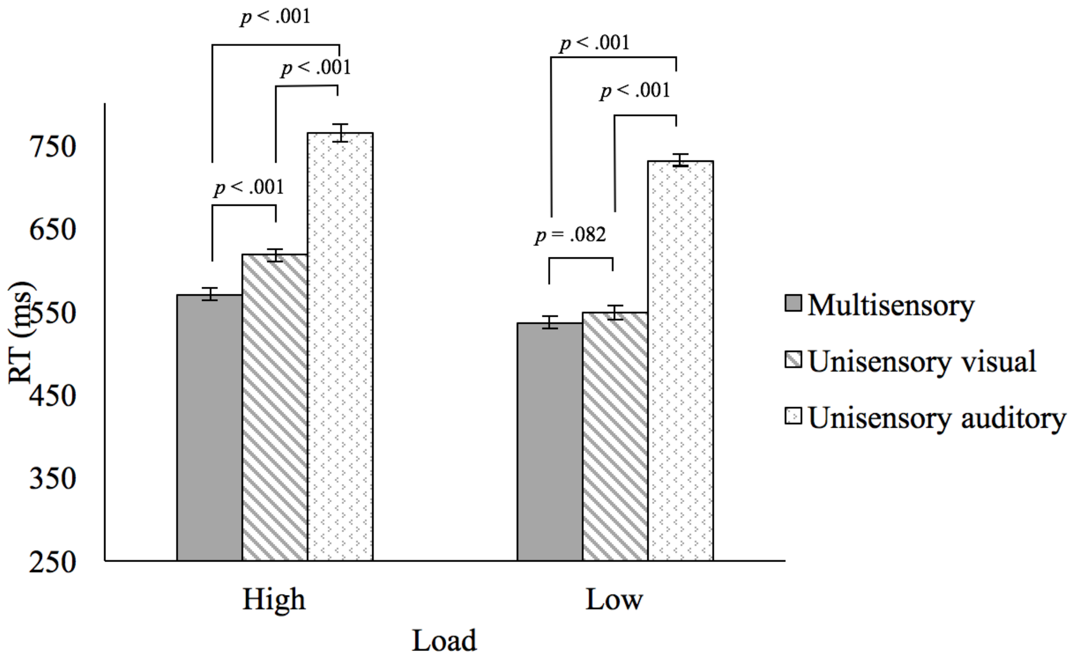

Figure 6: Reaction time (ms) of detection of peripheral targets in different modalities, as a function of load, in Experiment 3. Error bars show +/- 1 SEM with Cousineau-Morey correction (Cousineau, 2005; Morey, 2008)

5. Experiment 4: Salient Unisensory v. Multisensory Peripheral Distractors and High v. Low Load Central Task

Experiments 1-3 demonstrate facilitatory attentional capture by multisensory stimuli, in terms of both faster (Experiments 1-3) and more accurate (Experiments 2-3) detection of peripheral targets (relative to unisensory stimuli). This is consistent with the notion of multisensory stimuli having a special attentional status in terms of facilitated detection, although they do not appear to be entirely immune to the effects of perceptual load. Experiment 4 was designed to test whether attentional capture by multisensory stimuli would extend beyond facilitation effects, which necessarily involve the allocation of some top-down attention to the locations in which the multisensory stimuli appear. To test whether these effects extend to task irrelevant distractor stimuli, we adapted the protocol used in Experiment 3 to an ‘Irrelevant Distractor Task’ (Forster & Lavie, 2008), which has been previously established in the unisensory visual domain. The paradigm was similar to Experiment 3 with the exception that, as in the Irrelevant Distractor Task, participants were instructed to ignore the peripheral stimuli rather than respond to them. Now, capture by the peripheral targets should be inferred from their capacity to slow-down responses to the central task events (i.e. distractor interference). To maintain the low frequency of the distractors, which is necessary to observe a strong irrelevant distractor effect (Forster & Lavie, 2008b), whilst also maintaining an adequate number of trials in each condition, multisensory distractors were compared here with unisensory visual distractors (as used in the original version of the Irrelevant Distractor Task) [2].

5.1.Materials and Methods

5.1.1. Participants

52 participants (39 female) aged between 18 and 27 years (*M* = 20.26, *SD* = 1.94) were recruited at the University of Sussex. Seven participants were excluded for failing to reach an average of over 75% accuracy across the experiment. Participants were recruited until Bayes Factors for the main effect of load, and all planned distractor cost comparisons on reaction time reached sensitivity. Participants either gained course credits, or were paid, to take part. All participants reported normal or corrected-to-normal vision and hearing.

5.1.2. Stimuli and procedure

Stimuli and procedure are identical to Experiment 3, with the exception that participants were instructed to ignore anything in the periphery which may distract them from their task, that no auditory-only distractors were presented, and that there were two potential central targets (as in e.g. Forster & Lavie, 2008). Participants were instructed to search for either an X or an N in both high and low load conditions. On 16% of the trials, a distractor was presented, half being visual only and half being multisensory. They could not appear in the first three trials of each block. All distractors were presented to every participant, with the load, target

---

[2] Note that the fact that multisensory distractor stimuli do not, by their nature, require a response precluded a race model analysis in this experiment. In all of the present paper's previous experiments, visual stimuli produced greater capture effects than auditory stimuli and hence appeared the most competitive control condition. Had we found evidence of any multisensory enhancement of irrelevant distraction we would have proceeded to conduct a second experiment using unisensory auditory distractors – however, in the absence of any such effect this further experiment was not necessary (as any true multisensory benefit would be found in comparison to both visual and auditory distractors).

position, target identity, distractor side, distractor type and distractor identity, fully randomised.

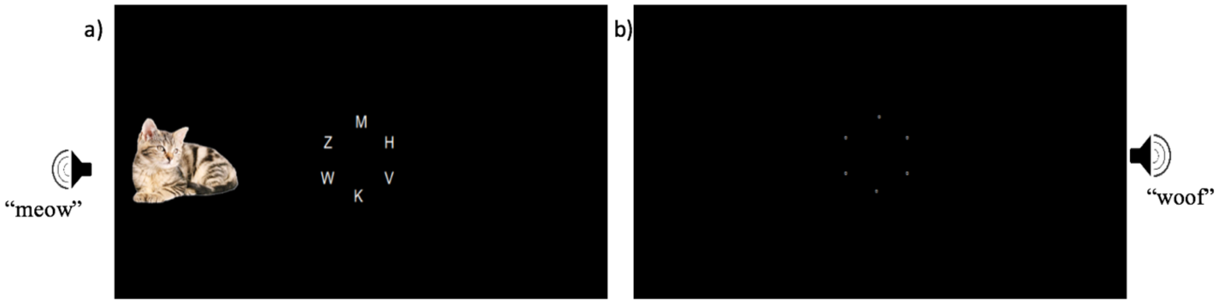

Figure 7: Example stimulus displays: a) high load, multisensory distractor, b) low load, auditory distractor

Participants completed three slow example trials, followed by 12 practice trials, for both high and low load. They then completed four blocks of 80 trials for each load, in the order ABBAABBA or BAABBAAB, which was counterbalanced between participants. Participants were instructed to respond as quickly as possible whilst still being accurate, and told to ignore anything else which may be presented to them other than the circle of letters. They had 2000ms make a response, a short beep indicated where this had been incorrect.

5.1.3. Bayesian analysis

Bayes factors were calculated using a half-normal distribution, with SDs based on Forster and Lavie's (2008) irrelevant distractor study using the same paradigm.

5.2.Results and Discussion

Mean RTs to correct responses and percentage error rates, as a function of distractor condition and load, are displayed in Table 2.

5.2.1. Reaction time (RT)

A 2 x 3 within-subject ANOVA with the factors of load (low, high) and distractor (multisensory, unisensory, no distractor), revealed a main effect of load, $F(1, 45) = 261.89$, $p < .001$, $\eta^2 = .86$, $B_{H(0,176)} = 4.90 \times 10^{171}$. RTs were slower in the high load than the low load condition, reflecting the increased demands of the high load task. There was no main effect of distractor ($p = .194$), however there was a significant interaction between load and distractor,

$F(1.63, 71.67) = 4.83$, $p = .016$, $\eta^2 = .10$. Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = .77$, $p = .004$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .81$). As can be seen in Figure 8, this interaction reflected that both multisensory and unisensory distractors slowed down RTs, relative to the no distractor condition, in the low load condition only.

	Distractor Condition		
	Multisensory	Visual	No Distractor
Low Load			
RT(ms)	535 (13)	530 (10)	512 (9)
% Error	9	9	9
High Load			
RT(ms)	792 (19)	783 (18)	795 (18)
% Error	25	23	21

Table 2: Mean RTs (SE in parentheses) and error rates (%) as a function of load and distractor type.

Planned comparisons revealed sensitive evidence for interference from both distractor types under low load: RT was significantly slower in the presence of a multisensory distractor than when no distractor was present, $t(44) = 3.36$, $p < .001$, $B_{H(0,60)} = 59.64$, and in the presence of a unisensory distractor than when no distractor was present, $t(44) = 4.25$, $p < .001$, $B_{H(0,60)} = 1185.32$. By contrast, under high load, comparison of RTs in the presence of either multisensory or visual only distractor compared with no distractor revealed sensitive null effects ($p = .623$, $B_{H(0,60)} = .12$ and $p = .932$, $B_{H(0,60)} = .05$ respectively). Critically, RTs in the presence of a multisensory distractor did not differ from RTs in the presence of a visual only distractor, for either high load ($p = .932$, $B_{H(0,60)} = .40$) or low load ($p = .228$ $B_{H(0,60)} = .27$).

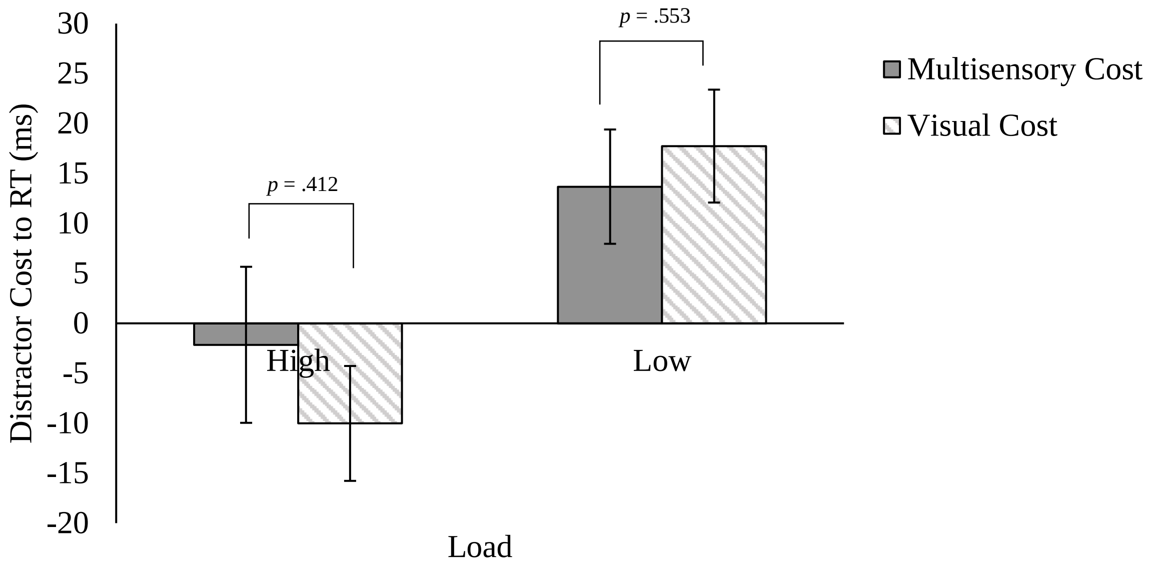

Figure 8: Cost to reaction time (ms) for detection of central target, due to the presence of a multisensory distract or visual distractor, as a function of load, error bars show +/- 1 SEM with Cousineau-Morey correction (Cousineau, 2005; Morey, 2008)

5.2.2. Error

A 2 x 3 within-subject ANOVA with the factors of load (low, high) and distractor type (multisensory, unisensory visual, no distractor), revealed a main effect of load, $F(1,44) = 78.96$, $p < .001$, $\eta^2 = .64$, $B_{H(0,4)} = 1.64\text{x}10^{34}$. Error rates were lower in the low load than the high load condition. There was no main effect of distractor type found ($p = .084$), or interaction between load and distractor type ($p = .163$).

The present results for visual distractors replicate previous findings using this paradigm (Forster & Lavie, 2008, 2014, 2016), as the effect of distractor presence on reaction time to the visual search task, present under low load conditions, was eliminated under high perceptual load. However, this pattern was seen not only for visual only distractors (the canonical effect) but also when the distractors were multisensory, demonstrating that their effects are not immune to perceptual load. Furthermore, unlike Experiments 1-3, this Experiment produced a striking absence of any evidence for enhanced attentional capture by multisensory stimuli above and beyond unisensory stimuli in any condition of load – these being sensitive null result under low load, falling just short of sensitivity under high load.

6. General Discussion

The present study sought to determine to what extent multisensory stimuli may be particularly effective at capturing attention, under both high and low perceptual load conditions. We have used, for the first time, established and controlled manipulations of perceptual load comparable to the ones traditionally used in unisensory perceptual load studies. The first key finding from this study is that across three experiments (Experiments 1-3), involving both high and low salience distractors, we demonstrate clear evidence of facilitatory attentional capture by multisensory events, in terms of both faster and more accurate detection compared to unisensory stimuli. This replicates a well-known multisensory advantage, previously reported in many studies and different paradigms (e.g. Colonius & Diedrich, 2002; Frassinetti, Bolognini & Làdavas 2002; Pérez-Bellido, Soto-Faraco & López-Moliner, 2013; Pannunzi et al., 2014). In contrast, however, we did not find any evidence of greater distractor interference by multisensory stimuli (Experiment 4). Note that Experiments 3 and 4 involved the same high salience peripheral stimuli - the key difference being whether participants were instructed to attend (and respond to) the peripheral stimuli (Experiment 3), or to ignore them (Experiment 4) – yet multisensory stimuli produced enhanced attentional capture in the former, but not latter, case. This is consistent with suggestions that multisensory integration is compromised when the stimuli involved are not attended (e.g., Alsius, Navarra, Campbell and Soto-Faraco, 2005; Talsma, Doty & Woldorff, 2007). In other words, multisensory integration, and hence enhanced attentional capture by multisensory events, may only occur for things which we are already looking out for.

The second key finding to emerge from the present study is that across all four experiments, and across two different manipulations of perceptual load, multisensory stimuli were not strictly ‘immune’ to perceptual load effects as has been previously proposed (Santangelo & Spence, 2007). Using a controlled, standard, manipulation of perceptual load, our Experiments 1-3 show that, similar to unisensory peripheral events, RTs to multisensory peripheral targets was slowed down when the central task was high perceptual load, and in Experiment 4 multisensory distractor costs were reduced under these conditions just like distractors costs associated to unisensory (visual) distractors. Hence, the processing of multisensory stimuli was modulated by load.

On the other hand, our findings are broadly compatible with the suggestion of Santangelo and Spence that multisensory stimuli might be particularly useful (e.g. as alerts) during high load tasks. Even under the most perceptually demanding conditions of our experiment, there was a detection time advantage towards multisensory stimuli. In fact, the multisensory capture effects observed in Experiments 1-3 were particularly pronounced in perceptually demanding situations, with significant violations of the race model found only under high perceptual load. While this could potentially imply greater resistance to perceptual load effects, we cannot rule out the alternative possibility that this could simply be a floor effect – that multisensory facilitation cannot decrease reaction time beyond a certain point which can already be achieved by unisensory stimuli under low load. In either case, from an applied perspective, multisensory stimuli may present greater advantages during demanding tasks. From a theoretical viewpoint, for now we conclude simply that multisensory stimuli do not appear to belong to the ‘special’ class of stimuli which are fully immune to the effects of perceptual load (e.g. human faces; Lavie, Ro & Russell, 2003).

The results of our research have not only theoretical implications, discussed above, but also practical implications for real-world scenarios. When driving a car, or focussing in a lecture, our results imply that an irrelevant multisensory event may be no more distracting than an already distracting unisensory one. On the other hand, if warning signals or alerts may come from a location the driver is monitoring already (e.g. a particular place on the dashboard), having this as a multisensory signal could mean a faster reaction time to detect it, compared to a unisensory one of equivalent strength. In addition, according to our results, detection may still be slower if driving through a busy town (high perceptual load) than down an empty lane (low perceptual load) for either unisensory or multisensory warning signals, although the multisensory advantage mentioned above would still be present in both circumstances.

A limitation of the present research is that our perceptual load task was always unisensory. An interesting question to consider in further research is whether multisensory capture would still be observed even if the load task itself, here the central stream monitoring, was multisensory. Research exploring this possibility is currently underway. A further fruitful direction for future research would be to identify the degree of task-relevance that is sufficient to allow multisensory enhancement of attentional capture – for example, is directing attention to the location of an item sufficient, or is it necessary to adopt an attentional setting for this item? This could be tested by adapting a task such as the singleton attentional capture task (e.g. Theeuwes, 1992), in which salient distractors appear as non-targets within the search array. Our findings could also be extended by testing whether our null findings concerning multisensory enhancement of irrelevant distraction could be replicated within other measures of distraction, for example temporal measures such as the attentional blink (Raymond, Shapiro & Arnell, 1992).

In conclusion, the present research points to a nuanced bidirectional relationship between multisensory integration and attention. On one hand, our results support the possibility that multisensory integration can, in certain contexts, enhance attention. On the other hand, our findings support suggestions that some degree of endogenous attention must be in place before integration (and hence any resulting attentional enhancement) may occur. When it does occur, multisensory enhancement of attention is further modulated by the availability of perceptual capacity, but may nevertheless be usefully exploited in applied contexts during demanding and undemanding conditions alike. As such our findings build on recent theoretical perspectives (e.g., Ten Oever et al., 2016, Hartcher-O’Brien, Soto-Faraco & Adam, 2017), by revealing a paradoxical interplay between integration and attention: multisensory processing may enhance attention, but only if you are already paying attention.

Conflict of Interest

The authors declare no conflicts of interest with respect to the authorship or the publication of this article.

References

Alsius, A., Möttönen, R., Sams, M. E., Soto-Faraco, S., & Tiippana, K. (2014). Effect of attentional load on audiovisual speech perception: evidence from ERPs. *Frontiers in psychology*, *5*, 727.

Alsius, A., Navarra, J., & Soto-Faraco, S. (2007). Attention to touch weakens audiovisual speech integration. Experimental Brain Research, 183(3), 399-404.

Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. Current Biology, 15(9), 839-843.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, *108*(25), 10367-10371.

Bahrami, B., Carmel, D., Walsh, V., Rees, G., & Lavie, N. (2007). Attentional load modulates subconscious orientation processing. *Journal of Vision*, *7*(9), 788-788.

Colonius, H., & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: a time-window-of-integration model. Journal of cognitive neuroscience, 16(6), 1000-1009.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method. *Tutorials in quantitative methods for psychology*, *1*(1), 42-45.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in psychology*, *5*, 781.

Fernández, L. M., Visser, M., Ventura-Campos, N., Ávila, C., & Soto-Faraco, S. (2015). Top-down attention regulates the neural expression of audiovisual integration. *NeuroImage*, *119*, 272-285.

Frassinetti, F., Bolognini, N., & Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental brain research*, *147*(3), 332-343.

Forster, S., & Lavie, N. (2007). High perceptual load makes everybody equal. *Psychological science*, *18*(5), 377-381.

Forster, S., & Lavie, N. (2008a). Failures to ignore entirely irrelevant distractors: the role of load. Journal of Experimental Psychology: Applied, 14(1), 73.

Forster, S., & Lavie, N. (2008b). Attentional capture by entirely irrelevant distractors. *Visual cognition*, *16*(2-3), 200-214.

Forster, S., & Lavie, N. (2014). Distracted by your mind? Individual differences in distractibility predict mind wandering. *Journal of experimental psychology: learning, memory, and cognition*, *40*(1), 251.

Forster, S., & Lavie, N. (2016). Establishing the attention-distractibility trait. *Psychological science*, *27*(2), 203-212.

Faul F, Erdfelder E, Buchner A, Lang AG (2009) Statistical power analyses using G* Power 31: Tests for correlation and regression analyses. Behav Res Methods 41:1149-1160.

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Attention, Perception, & Psychophysics*, *65*(7), 999-1010.

Hartcher-O'Brien, J., Soto-Faraco, S., & Adam, R. (2017). a matter of bottom-up or top-down processes: the role of attention in multisensory integration. *Frontiers in integrative neuroscience*, *11*, 5.

Ho, C., Reed, N., & Spence, C. (2007). Multisensory in-car warning signals for collision avoidance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *49*(6), 1107-1114.

Hughes, H. C., Reuter-Lorenz, P. A., Nozawa, G., & Fendrich, R. (1994). Visual-auditory interactions in sensorimotor processing: saccades versus manual responses. Journal of Experimental Psychology: Human Perception and Performance, 20(1), 131.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, *43*(4), 346-354.

Larson, G. E., & Merritt, C. R. (1991). Can accidents be predicted—An empirical test of the Cognitive Failures Questionnaire. *Applied Psychology—An International Review—Psychologie Appliquee—Revue Internationale*, *40*, 37–45.

Laurienti, P. J., Kraft, R. A., Maldjian, J. A., Burdette, J. H., & Wallace, M. T. (2004). Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental brain research*, *158*(4), 405-414.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human perception and performance*, *21*(3), 451.

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current directions in psychological science*, *19*(3), 143-148.

Lavie, N., & Cox, S. (1997). On the efficiency of visual selective attention: Efficient visual search leads to inefficient distractor rejection. Psychological Science, 8(5), 395-396.

Lavie, N., Ro, T., & Russell, C. (2003). The role of perceptual load in processing distractor faces. *Psychological science*, *14*(5), 510-515.

[dataset] Lunn, J., Sjoblom, A., Ward, J., Soto-Faraco, S. & Forster, S. (2018). Multisensory Enhancement of Attention Depends on Whether You Are Already Paying Attention, *Open Science Framework (OSF)*, Available from: https://osf.io/cvy8k/, DOI 10.17605/OSF.IO/CVY8K

Macdonald, J. S., & Lavie, N. (2008). Load induced blindness. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1078.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. Nature, 264(5588), 746−748.

Matusz, P. J., Broadbent, H., Ferrari, J., Forrest, B., Merkley, R., & Scerif, G. (2015). Multi-modal distraction: Insights from children's limited attention. *Cognition*, *136*, 156-165.

Matusz, P. J., & Eimer, M. (2011). Multisensory enhancement of attentional capture in visual search. *Psychonomic bulletin & review*, *18*(5), 904.

Meredith, M. A., & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of neurophysiology*, *56*(3), 640-662.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive psychology*, *14*(2), 247-279.

Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics*, *40*(5), 331-343.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. Cognitive brain research, 14(1), 115-128.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *reason*, *4*(2), 61-64.

Murphy, G., & Greene, C. M. (2015). High perceptual load causes inattentional blindness and deafness in drivers. *Visual Cognition*, *23*(7), 810-814.

Murphy, G., & Greene, C. M. (2016). Perceptual load induces inattentional blindness in drivers. *Applied Cognitive Psychology*, *30*(3), 479-483.

Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., ... & Foxe, J. J. (2004). Grabbing your ear: rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. Cerebral cortex, 15(7), 963-974.

Otto, T. U., & Mamassian, P. (2012). Noise and correlations in parallel perceptual decision making. *Current Biology*, *22*(15), 1391-1396.

Pannunzi, M., Pérez-Bellido, A., Pereda-Baños, A., López-Moliner, J., Deco, G., & Soto-Faraco, S. (2014). Deconstructing multisensory enhancement in detection. Journal of neurophysiology, 113(6), 1800-1818.

Pápai, M. S., & Soto-Faraco, S. (2017). Sounds can boost the awareness of visual events through attention without cross-modal integration. Scientific Reports, 7, 41684.

Pérez-Bellido, A., Soto-Faraco, S., & López-Moliner, J. (2012). Sound-driven enhancement of vision: disentangling detection-level from decision-level contributions. *Journal of Neurophysiology*, *109*(4), 1065-1077.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, *32*(1), 3-25.

Purkis, H. M., Lester, K. J., & Field, A. P. (2011). But what about the Empress of Racnoss? The allocation of attention to spiders and Doctor Who in a visual search task is predicted by fear and expertise. *Emotion*, *11*(6), 1484.

Raveh, D., & Lavie, N. (2015). Load-induced inattentional deafness. *Attention, Perception, & Psychophysics*, *77*(2), 483-492.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink?. *Journal of experimental psychology: Human perception and performance*, *18*(3), 849.

Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin & Review*, *21*(2), 301-308.

Santangelo, V., Botta, F., Lupiáñez, J., & Spence, C. (2011). The time course of attentional capture under dual-task conditions. *Attention, perception, & psychophysics*, *73*(1), 15-23.

Santangelo, V., Ho, C., & Spence, C. (2008). Capturing spatial attention with multisensory cues. *Psychonomic Bulletin & Review*, *15*(2), 398-403.

Santangelo, V., & Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(6), 1311.

Schwartz, S., Vuilleumier, P., Hutton, C., Maravita, A., Dolan, R. J., & Driver, J. (2004). Attentional load and sensory competition in human vision: modulation of fMRI responses by load at fixation during task-irrelevant stimulation in the peripheral visual field. *Cerebral cortex*, *15*(6), 770-786.

Senkowski, D., Talsma, D., Herrmann, C. S., & Woldorff, M. G. (2005). Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. Experimental Brain Research, 166(3-4), 411-426.

Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration?. *Cerebral cortex*, *17*(3), 679-690.

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in cognitive sciences*, *14*(9), 400-410.

Talsma, D., & Woldorff, M. G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *Journal of cognitive neuroscience*, *17*(7), 1098-1114.

Ten Oever, S., Romei, V., van Atteveldt, N., Soto-Faraco, S., Murray, M. M., & Matusz, P. J. (2016). The COGs (context, object, and goals) in multisensory processing. *Experimental brain research*, *234*(5), 1307-1323.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Attention, Perception, & Psychophysics*, *51*(6), 599-606.

Ulrich, R., Miller, J., & Schröter, H. (2007). Testing the race model inequality: An algorithm and computer programs. *Behavior Research Methods*, *39*(2), 291-302.

Van der Burg, E., Olivers, C. N., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1053.

Wagenmakers, E. J., Verhagen, A. J., Ly, A., Matzke, D., Steingroever, H., Rouder, J. N., & Morey, R. D. (2015). The need for Bayesian hypothesis testing in psychological science. *Psychological science under scrutiny: Recent challenges and proposed solutions*.